Consensus over a Random Network Generated by i.i.d. Stochastic Matrices

Qingshuo Song, Guanrong Chen, Daniel W. C. Ho

Abstract

Our goal is to find a necessary and sufficient condition on the consensus over a random network, generated by i.i.d. stochastic matrices. We show that the consensus problem in three different convergence modes (almost surely, in probability, and in L^1) are equivalent, thus have the same necessary and sufficient condition. We obtain the necessary and sufficient condition through the stability in a projected subspace.

Keywords and Phrases. Consensus, stability, random network, stochastic matrix.

I. Introduction

We consider a stochastic linear difference equation,

$$X(t) = A(t)X(t-1), \quad t = 1, 2, \dots$$

where the states $\{X(t)\}$ is an $\mathbb{R}^N$-valued sequence, and $\{A(t)\}$ is a sequence of i.i.d. (independent and identically distributed) right stochastic matrices (non-negative matrix with each row summing to 1). The system is said to reach consensus if, for any initial state, $\max_{1\leq i,j\leq N}|X_i(t)-X_j(t)|$ converges to zero as $t\to\infty$ in an appropriate sense. Since $X(t)$ is random, there are different modes of consensus: Almost surely consensus, in probability consensus, and L^1 consensus.

Department of Mathematics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, song.qingshuo@cityu.edu.hk.

Department of Electronic Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, gchen@ee.cityu.edu.hk.

Department of Mathematics, City, University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, madaniel@cityu.edu.hk.

Consensus problem over a difference equation has wide applications in random network theory (e.g., [3], [4] and the references therein). [3] studies the consensus in probability. [4] establish an elegant necessary and sufficient condition for almost surely consensus by investigating the ergodicity of a random matrix sequence:

$$|\lambda_2(\mathbb{E}[A(1)])| < 1 \tag{I.1}$$

where $\mathbb{E}[\cdot]$ is the expectation operator and $\lambda_2(\cdot)$ is the second largest eigenvalue (in absolute value) of the argument matrix. Note that, almost surely consensus implies in probability consensus, and thus (I.1) is obviously a sufficient condition for in probability consensus.

In this work, looking further into the specific nature of $\{X(t)\}$, we show that the consensus in all three modes are actually indifferent, hence (I.1) gives necessary and sufficient condition for consensus in all three modes. In addition, by using a completely different methodology in contrast to [4], our result applies to a more general setting: their restriction on the space of stochastic matrices with strictly positive diagonal entries can be relaxed (see Remark II.1).

The main ingredient of our work is that, based on the observation of a relation between consensus and stability, the original consensus problem on a sequence is reduced to the stability problem on a projected sequence in a subspace. As a result, we can focus our study on the eigenspace structure of the projection operator. As a by-product, we offer a simple proof of consensus to deterministic linear networks.

The rest of the paper is arranged as follows: We start with the problem formulation in section 2, where a crucial result on the relation between consensus and stability is presented. In section 3, a simple proof of consensus on a deterministic sequence is provided, which can be read independently for readers only interested in the deterministic case. The main result, the necessary and sufficient condition for consensus of a random network, is established in section 4. Finally, we conclude our investigation in section 5.

II. Problem formulation

In the first subsection, a crucial result (Theorem II.1) on the equivalence of consensus and stability in a subspace will be presented under a general setup of consensus problem. This theorem can be applied to very general cases setup, including nonlinear and random sequences, and plays an important role throughout the paper. In the second subsection, the main consensus problem is formulated using a linear stochastic difference equation.

Before proceeding, let us recall some standard notations:

1) In (column) vector space $\mathbb{R}^N$, x_i is the ith coordinate of vector $x \in \mathbb{R}^N$; l^p-norm is $\|x\|_p = \sum_{i=1}^N |x_i|^p$, $\forall 1 \le p \le \infty$; x^T denotes the transpose of x.
2) In square real matrix space $\mathbb{R}^{N\times N}$, I is the identity matrix; for all $A \in \mathbb{R}^{N\times N}$, $\|A\|_p = \max_{\|x\|=1} \|Ax\|_p$ for all $1 \le p \le \infty$; the eigenvalues will be arranged in order of $|\lambda_1(A)| \ge |\lambda_2(A)| \ge \cdots \ge |\lambda_N(A)|$; the spectral radius refers to $\rho(A) = |\lambda_1(A)|$.
3) $\|\cdot\|$ is used in the formula if it is valid for all l^p-norms.
4) Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we denote by $\mathbb{E} = \mathbb{E}^{\mathbb{P}}$ the expectation under $\mathbb{P}$. L^p refers to $L^p(\Omega, \mathcal{F}, \mathbb{P})$: for random vector $Y : \Omega \to \mathbb{R}^N$, the L^p-norm is $\|Y\|_{L^p} = (\int_\Omega \|Y(\omega)\|_2 \mathbb{P}(d\omega))^{1/p} = (\mathbb{E}[(\|Y\|_2)^p])^{1/p}$.

A. *A general consensus problem*

Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a filtered probability space, where $\mathbb{F} = \{\mathcal{F}_t : t = 0, 1, 2, \ldots\}$ is a sequence of increasing σ-algebras with $\mathcal{F}_\infty \subset \mathcal{F}$. We consider an $\mathbb{F}$-adapted sequence $\{X(t)\}$ taking values in $\mathbb{R}^N$. In other words, $X(t)$ is a measurable mapping from $(\Omega, \mathcal{F}_t) \to (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, where $\mathcal{B}(\mathbb{R}^N)$ is the Borel σ-algebra on $\mathbb{R}^N$. Such a sequence includes the general form of

$$X(t) = f_t(X(t-1), \ldots, X(1))$$

for some measurable function f_t, and emphasizes its independence of future events First, we start from the precise definition of consensus on random sequence in three different modes. As usual, $X(t, \omega)$ will be used instead of $X(t)$ when we need to emphasize its dependence on a sample path $\omega \in \Omega$.

Definition II.1 (Consensus of a sequence)**.** *Let $\{X(t)\}$ be an $\mathbb{F}$-adapted $\mathbb{R}^N$-valued random sequence. $\{X(t)\}$ is said to reach consensus*

1) *in probability, if*

$$\lim_{t\to\infty} \mathbb{P}\Big\{\omega \in \Omega : \max_{1\le i,j\le N} |X_i(t,\omega) - X_j(t,\omega)| > \varepsilon\Big\} = 0, \ \forall \varepsilon > 0.$$

2) *almost surely* (*with probability* 1)*, if*

$$\mathbb{P}\Big\{\omega \in \Omega : \lim_{t\to\infty} \max_{1\le i,j\le N} |X_i(t,\omega) - X_j(t,\omega)| = 0\Big\} = 1.$$

3) *in* L^p $(p \geq 1)$, *if*

$$\lim_{t\to\infty} \mathbb{E}\Big[\max_{1\leq i,j\leq N} |X_i(t) - X_j(t)|^p\Big] = 0.$$

Definition II.2 (Stability of a sequence)**.** $\{X(t)\}$ *is said to be stable* $($*at zero*$)$

1) *in probability, if*

$$\lim_{t\to\infty} \mathbb{P}\Big\{\omega \in \Omega : \|X(t,\omega)\| > \varepsilon\Big\} = 0, \ \forall\varepsilon > 0.$$

2) *almost surely, if*

$$\mathbb{P}\Big\{\omega \in \Omega : \lim_{t\to\infty} \|X(t,\omega)\| = 0\Big\} = 1.$$

3) *in* L^p $(p \geq 1)$, *if*

$$\lim_{t\to\infty} \mathbb{E}\Big[\|X(t,\omega)\|^p\Big] = 0.$$

Define a subspace of $\mathbb{R}^N$ by $\mathbb{R}_0 = \{x \in \mathbb{R}^N : x_1 = x_2 = \ldots = x_N\}$. Let Π be a projection operator on $\mathbb{R}_0$, i.e.

$$\Pi x = \langle x, v_0\rangle v_0, \quad \forall x \in \mathbb{R}_0,$$

where $\langle\cdot,\cdot\rangle$ is inner product, $v_0 \in \mathbb{R}_0$ is an l^2-norm unit vector. We thus have the orthogonal projection $\Pi^\perp : \mathbb{R}^N \to \mathbb{R}^\perp$ by $\Pi^\perp = I - \Pi$, so that the orthogonal decomposition is valid

$$x = \Pi x + \Pi^\perp x, \ \forall x \in \mathbb{R}^N. \tag{II.1}$$

The following theorem shows that the consensus of a sequence in $\mathbb{R}^N$ is equivalent to the stability of the sequence projected on the subspace $\mathbb{R}_0^\perp$.

Theorem II.1. $\{X(t)\}$ *reaches consensus almost surely* $($*respectively, in probability, or in* $L^p)$ *if and only if* $\{\Pi^\perp X(t)\}$ *is stable almost surely* $($*respectively, in probability, or in* $L^p)$.

Proof: We will show the equivalence of stability and consensus in the sense of almost surely. The equivalence in probability and in L^p can be similarly proved.

($\Longrightarrow$) Suppose $\{X(t)\}$ reaches consensus almost surely. Define $Y(t) \in \mathbb{R}_0$ be a vector with all entries equal to the value of first coordinate of $X(t)$, i.e. $Y(t) = (X_1(t), X_1(t), \ldots, X_1(t))^T$. Since

$$\begin{aligned} \|\Pi^\perp X(t)\|_\infty &= \min\nolimits_{y\in\mathbb{R}_0} \|X(t) - y\|_\infty \\ &\leq \|X(t) - Y(t)\|_\infty \\ &= \max\nolimits_{1\leq i\leq N} |X_i(t) - X_1(t)| \\ &\leq \max\nolimits_{1\leq i,j\leq N} |X_i(t) - X_j(t)| \to 0 \end{aligned}$$

as $t \to \infty$ almost surely. Therefore, $\{\Pi^\perp X(t)\}$ is stable almost surely.

($\Longleftarrow$) Suppose $\{\Pi^\perp X(t)\}$ is stable almost surely. Let $(\Pi X)_i(t)$ is the ith coordinate of vector $\Pi X(t)$. Note that, since $\Pi X(t) \in \mathbb{R}_0$, we have all coordinates with the same value, that is, $(\Pi X)_i(t) = (\Pi X)_j(t), \forall i, j$. Therefore, by triangle inequality,

$$\begin{aligned}
\max_{i,j} |X_i(t) - X_j(t)| &\leq \max_{i,j}(|X_i(t) - (\Pi X)_i(t)| + |(\Pi X)_j(t) - X_j(t)|) \\
&\leq \max_i |X_i(t) - (\Pi X)_i(t)| + \max_j |(\Pi X)_j(t) - X_j(t)| \\
&\leq 2\|X(t) - \Pi X(t)\|_\infty \\
&= 2\|\Pi^\perp X(t)\|_\infty \to 0
\end{aligned}$$

as $t \to \infty$ almost surely. Therefore, $\{X(t)\}$ reaches consensus almost surely.

■

B. Consensus problem over linear random network

We consider a similar setting as [4]. Let the space of $N \times N$ stochastic matrices be

$$S_N = \left\{ A = (a_{ij})_{N\times N} : a_{ij} \geq 0,\ \sum_{j=1}^{N} a_{ij} = 1, \forall i, j \right\} \tag{II.2}$$

and $\mathcal{B}(S_N)$ be the Borel σ-algebra on S_N. Let μ be a given probability distribution on $(S_N, \mathcal{B}(S_N))$, and $\{A(t)\}$ be an S_N-valued i.i.d. sequence with distribution μ. Let $\Omega = (S_N)^\infty$, $\mathbb{P} = \mu \times \mu \times \cdots$, and $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_t = \sigma(A(1), A(2), \ldots, A(t))$ for $t \geq 1$, $\mathbb{F} = \cup_{t=0}^\infty \mathcal{F}_t$.

Now, we consider a random sequence $\{X(t)\}$ given by

$$X(t) = A(t)X(t-1),\ \forall t \in \mathbb{N};\ X(0) = x. \tag{II.3}$$

Then, $\{X(t)\}$ is an $\mathbb{F}$-adapted sequence in the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. Observe that the distribution of $X(t)$ is determined by the initial state $X(0) = x$ and distribution μ. Sometimes, we write $X^x(t)$ to emphasize the initial state $X(0) = x$ in the context.

Definition II.3 (Consensus and stability of a distribution)**.** *A distribution μ is said to reach consensus almost surely* (*respectively, in probability, or in L^p*), *if $\{X^x(t)\}$ of* (II.3) *generated by the distribution μ reaches consensus almost surely* (*respectively, in probability, or in L^p*) *for all initial states $x \in \mathbb{R}^N$.*

Similar to Definition II.3, one can define stability for the sequence (II.3) generated by the distribution μ.

Remark II.1. [4] had a different problem formulation in that the space S_N of (II.2) was replaced by a smaller space

$$\hat{S}_N = \{A \in S_N : \text{ all diagonal entries are strictly positive } \}.$$

Indeed, such a restriction is crucial in the proof of [4, Theorem 3] to utilize the [1, Perron-Frobenius theorem] on primitive matrix. In our work, the diagonal entries can be zero, which therefore covers the results of [4] as a special case with $\mu(S_N \setminus \hat{S}_N) = 0$. $\square$

Next, our goal is to find a necessary and sufficient condition for the consensus of the distribution μ.

III. Necessary and sufficient condition for a deterministic sequence

A deterministic system can be treated as a special case of a random system in the following sense. Let the probability distribution μ on S_N satisfy $\mu(\{A\}) = 1$ for some stochastic matrix $A \in S_N$. Then, $A(t) = A$ for all $t = 1, 2, \ldots$, and the sequence $\{X(t)\}$ of (II.3) becomes deterministic, and in this case we have

$$X(t) = A^t x, \ \forall t = 0, 1, 2, \ldots$$

For convenience, we say A reaches consensus if $\{X(t) = A^t x\}$ reaches consensus for all initial states $x \in \mathbb{R}^N$. Note that, the deterministic consensus is indifferent to all three modes of consensus, since the sample space Ω can be treated as a singleton $\{A\} \times \{A\} \times \cdots$. Thus, this definition is consistent with Definition II.3 on consensus (in all three modes) of distribution μ of the form $\mu(\{A\}) = 1$.

The main idea in this section is that, thanks to Theorem II.1, it is equivalent to find a sufficient and necessary condition of the stability of $\{\Pi^{\perp} X(t)\}$, which turns out to be a sequence generated by the projection matrix $\Pi^{\perp} A$. We will show that $\Pi^{\perp} A$ is stable if and only if $\rho(\Pi^{\perp} A) < 1$ by Proposition III.2. Together with the fact that $\rho(\Pi^{\perp} A) = |\lambda_2(A)|$, by Proposition III.1, we will obtain the desired necessary and sufficient condition.

To proceed with the consensus on the deterministic sequence $X(t)$ generated by the stochastic matrix A, we first recall some properties of stochastic matrices. Since each row sum of a stochastic matrix is equal to 1, its largest eigenvalue is $\rho(A) = \lambda_1(A) = 1$, i.e. $Ax = x$

for all $x \in \mathbb{R}_0$. Also, we have

$$\|A^t x\|_\infty \le \|x\|_\infty, \ \forall x \in \mathbb{R}^N, t \in \mathbb{N}. \tag{III.1}$$

In addition, we have the following useful results:

Proposition III.1. *Let $A \in S_N$ be a stochastic matrix. Then*

1) *A has a Jordan canonical form of*

$$\Lambda = \begin{bmatrix} 1 & 0_{1\times(N-1)} \\ 0_{(N-1)\times 1} & \Lambda_{22} \end{bmatrix}, \tag{III.2}$$

where $0_{m\times n}$ is $m \times n$ matrix with each entry being zero, and Λ_{22} is sub-matrix of Jordan form.

2) *The linear operator $\Pi^\perp$ defined in* (II.1) *satisfies*

$$\Pi^\perp A = \Pi^\perp A \Pi^\perp \tag{III.3}$$

3) *$\Pi^\perp A$, as a matrix, has a Jordan form of $\Lambda_0 = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_{22} \end{bmatrix}$ with Λ_{22} defined in* (III.2). *In particular, $\rho(\Pi^\perp A) = |\lambda_2(A)|$.*

Proof:

1) Let v_0 be the unit vector in the subspace $\mathbb{R}_0$. Note that $v_0 \in \mathbb{R}_0$ is an eigenvector of A associated with the eigenvalue $\lambda_1(A) = 1$, i.e. $(A - I)v_0 = 0$. To prove the first claim, we only need to show that the Jordan block corresponding to $\lambda_1(A) = 1$ is simple. If not, there exists $v_1 \notin \mathbb{R}_0$ associated with the Jordan block $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, satisfying

$$(A - I)v_1 = v_0. \tag{III.4}$$

By induction,

$$A^t v_1 = t \cdot v_0 + v_1, \ \forall t \in \mathbb{N}.$$

This implies that $\|A^t v_1\|_\infty \to \infty$ as $t \to \infty$, which leads to a contradiction to (III.1).

2) One can prove (III.3) as follows: $\forall x \in \mathbb{R}^N$

$$\Pi^\perp A x = \Pi^\perp A(\Pi x + \Pi^\perp x) = \Pi^\perp A \Pi x + \Pi^\perp A \Pi^\perp x = \Pi^\perp \Pi x + \Pi^\perp A \Pi^\perp x = \Pi^\perp A \Pi^\perp x.$$

3) If $\{x_1, \dots x_m\}$ are generalized eigenvectors of A associated with some eigenvalue λ in the Jordan block in Λ_{22}, satisfying

$$(A - \lambda I)x_i = x_{i-1}, \text{ for } i = 1, 2, \dots m, \quad x_0 = 0,$$

then, by the facts $x_1 \notin \mathbb{R}_0$ and (III.3),

$$\begin{aligned}(\Pi^\perp A - \lambda I)(\Pi^\perp x_i) &= \Pi^\perp A \Pi^\perp x_i - \lambda \Pi^\perp x_i \\ &= \Pi^\perp A x_i - \lambda \Pi^\perp x_i \\ &= \Pi^\perp (A - \lambda I) x_i \\ &= \Pi^\perp x_{i-1}.\end{aligned}$$

In other words, since $\Pi^\perp x_1 \neq 0$, $\Pi^\perp A$ with $\{\Pi^\perp x_1, \dots, \Pi^\perp x_m\}$ preserves the structure of the eigenspace associated with matrix A corresponding to the eigenvalue λ in Jordan block Λ_{22}. Also, since λ is arbitrary eignevalue in the Jordan block Λ_{22}, together with $\Pi^\perp A v_0 = 0$, we conclude $\Pi^\perp A$ has a Jordan form of $\Lambda_0 = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_{22} \end{bmatrix}$. Finally, we have

$$\rho(\Pi^\perp A) = \rho(\Lambda_0) = \rho(\Lambda_{22}) = |\lambda_2(A)|.$$

■

Next, we present a necessary and sufficient condition for stability.

Proposition III.2. $A \in \mathbb{R}^{N \times N}$ *is stable if and only if* $\rho(A) < 1$.

Proof: One can use the fact $\lim_{t \to \infty} \|A^t\|^{1/t} = \rho(A)$ to complete the proof. ■

Thanks to Proposition III.1 and Proposition III.2, we are now ready to obtain a necessary and sufficient condition for the consensus of a deterministic sequence.

Theorem III.1 (Necessary and sufficient condition in deterministic case)**.** $A \in S_N$ *reaches consensus if and only if* $|\lambda_2(A)| < 1$.

Proof: By Theorem II.1, A reaches consensus if and only if $\{\Pi^\perp X(t)\}$ is stable. Note that, by (III.3), for any initial state $x \in \mathbb{R}^N$

$$\begin{aligned}\Pi^\perp X(t) &= \Pi^\perp A X(t-1) = (\Pi^\perp A) \Pi^\perp X(t-1) \\ &= \cdots = (\Pi^\perp A)^t \Pi^\perp x = (\Pi^\perp A)^t x.\end{aligned} \tag{III.5}$$

Thus, $\{\Pi^{\perp}X(t)\}$ is a sequence generated by $\Pi^{\perp}A$. By Proposition III.2, $\{\Pi^{\perp}X(t)\}$ is stable if and only if $\rho(\Pi^{\perp}A) < 1$. Observe that, by Proposition III.1, $\rho(\Pi^{\perp}A) = |\lambda_2(A)|$. This completes the proof. ■

IV. Necessary and sufficient condition for a stochastic sequence

In this section, we return to the stochastic sequence $\{X(t)\}$ defined in (II.3) generated by distribution μ, and study a necessary and sufficient condition for its consensus. First, by studying the fine structure of the random sequence generated by i.i.d. stochastic matrices, we show that consensus in three different modes classified by Definition II.1 are in fact equivalent to each other. Thus, we can only work on the almost surely consensus.

A. Equivalence of consensus in three modes

Before we proceed with the equivalence of consensus in three modes, we briefly recall some relations between convergence of random variables in three modes, and we refer to [2] for more detail. Consider a sequence of random variables $\{a_n, n = 1, 2, \ldots\}$ and a random variable $a \geq 0$. Both almost surely convergence and L^1 convergence imply in probability convergence, i.e. $a_n \to a$ almost surely implies $a_n \to a$ in probability; $a_n \to a$ in L^1 implies $a_n \to a$ in probability. However, the reverse directions need further conditions in general. $a_n \to a$ in probability together with $|a_n| \leq |b|$ almost surely for some $b \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ implies $a_n \to a$ in L^1 by the dominated convergence theorem; $a_n \to a$ in L^1 and $0 \leq a_{n+1} \leq a_n$ almost surely together implies $a_n \to a$ almost surely by the monotone convergence theorem.

Lemma IV.1. *Consider the sequence $\{X(t)\}$ defined in* (II.3) *generated by distribution μ. Given $X(0) = x$, the following statements on stability of $\{\Pi^{\perp}X(t)\}$ are equivalent:*

1) $\{\Pi^{\perp}X(t)\}$ *is stable in probability.*
2) $\{\Pi^{\perp}X(t)\}$ *is stable in L^1.*
3) $\{\Pi^{\perp}X(t)\}$ *is stable almost surely.*

Proof: Observe that, by (III.3), $\{\Pi^{\perp}X(t)\}$ is a sequence generated by random matrix $\Pi^{\perp}A(t)$, i.e.

$$\Pi^{\perp}X(t) = \Pi^{\perp}A(t)\ \Pi^{\perp}X(t-1). \tag{IV.1}$$

In the following, we prove the equivalence by showing: (1) implies (2), (2) implies (3), (3) implies (1), respectively.

1) If the sequence $\{\Pi^{\perp}X(t)\}$ is stable in probability, then $\|\Pi^{\perp}X(t)\|_{\infty} \to 0$ in probability. Together with the uniform boundedness $\|\Pi^{\perp}X(t)\|_{\infty} \le \|x\|_{\infty}$, the dominated convergence theorem implies that $\|\Pi^{\perp}X(t)\|_{\infty} \to 0$ in L^1. Thus, the sequence $\{\Pi^{\perp}X(t)\}$ is stable in L^1.
2) If the sequence $\{\Pi^{\perp}X(t)\}$ is stable in L^1, then $\|\Pi^{\perp}X(t)\|_{\infty} \to 0$ in L^1. In addition, one can show the monotonicity of $\|\Pi^{\perp}X(t)\|_{\infty} \le \|\Pi^{\perp}X(t-1)\|_{\infty}$, by observing

$$\|\Pi^{\perp}X(t)\|_{\infty} = \|\Pi^{\perp}A(t)X(t-1)\|_{\infty} = \|\Pi^{\perp}A(t)\Pi^{\perp}X(t-1)\|_{\infty} \le \|\Pi^{\perp}X(t-1)\|_{\infty}. \tag{IV.2}$$

   By the monotone convergence theorem, $\|\Pi^{\perp}X(t)\|_{\infty} \to 0$ almost surely.
3) It is well known that almost surely convergence implies convergence in probability.

■

The next theorem about equivalent consensus in three modes is a main result in our paper.

Theorem IV.1. *Consider the sequence $\{X(t)\}$ defined in* (II.3) *generated by distribution μ. The following statements on consensus are equivalent:*

1) *Distribution μ reaches consensus in probability.*
2) *Distribution μ reaches consensus in L^1.*
3) *Distribution μ reaches consensus almost surely.*

Proof: It follows from Theorem II.1 and Lemma IV.1. ■

B. Necessary and sufficient condition for the random case

Thanks to Theorem IV.1, our work is now reduced to finding a necessary and sufficient condition for consensus in any one of the three modes. Below, we say the distribution μ reaches consensus without specifying a convergence mode.

Recall from Definition II.3 that, a distribution μ reaches consensus if the generated sequence $\{X^x(t)\}$ reaches consensus for all initial states $x \in \mathbb{R}^N$. The next proposition shows that it is sufficient to check the consensus of $\{X^x(t)\}$ only for all $x \in (\mathbb{R}^+)^N$ to guarantee the consensus of a distribution μ.

Proposition IV.1. *μ reaches consensus if and only if $X^x(t)$ defined in* (II.3) *reaches consensus for all $x \in (\mathbb{R}^+)^N$.*

Proof: Observe that $X^{x+c}(t) = X^x(t) + c$ for all $c \in \mathbb{R}_0$. Hence,

$$\max_{1\le i,j\le N} |X_i^x(t,\omega) - X_j^x(t,\omega)| = \max_{1\le i,j\le N} |X_i^{x+c}(t,\omega) - X_j^{x+c}(t,\omega)|, \ \forall c \in \mathbb{R}_0.$$

In other words, to consider consensus of $\{X^x(t)\}$ for some $x \notin (\mathbb{R}^+)^N$, one can always investigate the consensus of $X^{x+c}(t)$ equivalently, by taking $c = (\|x\|_\infty, \ldots, \|x\|_\infty)^T \in \mathbb{R}_0$. Note $x + c \in (\mathbb{R}^+)^N$, hence the result holds. ■

Next, we review some useful properties of the expectation operator $\mathbb{E}$. First, the expectation operator $\mathbb{E}$ is commutative with any deterministic matrix A, i.e.,

$$A\mathbb{E}[Y] = \mathbb{E}[AY], \ \forall \mathcal{F}\text{-measurable } Y : \Omega \to \mathbb{R}^N. \tag{IV.3}$$

In particular, by taking $A = \Pi^\perp$, we have $\Pi^\perp \mathbb{E} = \mathbb{E}\Pi^\perp$. Furthermore, for an arbitrary random matrix A, if a random vector $Y : \Omega \to \mathbb{R}^N$ is independent of A, then

$$\mathbb{E}[AY] = \mathbb{E}[A]\mathbb{E}[Y]. \tag{IV.4}$$

Finally, note that the deterministic sequence $\{\mathbb{E}[X(t)]\}$ is actually a sequence generated by the deterministic matrix $\mathbb{E}[A]$, since by (IV.4)

$$\begin{aligned} \mathbb{E}[X(t)] &= \mathbb{E}[A(t)X(t-1)] \\ &= \mathbb{E}[A(t)]\mathbb{E}[X(t-1)] \\ &= \mathbb{E}[A(1)] \cdot \mathbb{E}[X(t-1)]. \end{aligned} \tag{IV.5}$$

Theorem IV.2 (Necessary and sufficient condition for consensus)**.** *Consider $\{X(t)\}$ defined in* (II.3) *generated by distribution μ. μ reaches consensus almost surely (also, in probability, and in L^1) if and only if $\lambda_2(\mathbb{E}^\mu[A(1)]) < 1$.*

Proof: By saying that μ reaches consensus, we mean μ reaches consensus in any of three modes due to Theorem IV.1.

($\Longrightarrow$) If μ reaches consensus, then Theorem II.1 implies $\Pi^\perp X(t) \to 0$ in L^1 for all initial states $X(0) = x$, hence $\mathbb{E}[\|\Pi^\perp X(t)\|_\infty] \to 0$.

Next, by (IV.3) and Jensen's inequality, we have

$$\|\Pi^\perp \mathbb{E}[X(t)]\|_\infty = \|\mathbb{E}[\Pi^\perp X(t)]\|_\infty \le \mathbb{E}[\|\Pi^\perp X(t)\|_\infty] \to 0$$

This implies that the deterministic sequence $\{\Pi^{\perp}\mathbb{E}[X(t)]\}$ is stable. Then, using (IV.5) and (III.3), we have

$$\Pi^{\perp}\mathbb{E}[X(t)] = \Pi^{\perp}\mathbb{E}[A(1)] \cdot \mathbb{E}[X(t-1)] = \Pi^{\perp}\mathbb{E}[A(1)] \cdot \Pi^{\perp}\mathbb{E}[X(t-1)]$$

In other words, $\{\Pi^{\perp}\mathbb{E}[X(t)]\}$ is a deterministic sequence generated by matrix $\Pi^{\perp}\mathbb{E}[A(1)]$. Thus, by Proposition III.1 and Proposition III.2, $\rho(\Pi^{\perp}\mathbb{E}[A(1)]) = |\lambda_2(\mathbb{E}[A(1)])| < 1$.

($\Longleftarrow$) It follows from (IV.5) that the deterministic sequence $\{\mathbb{E}[X(t)]\}$ is generated by matrix $\mathbb{E}[A(1)]$. If $|\lambda_2(\mathbb{E}[A(1)])| < 1$, then by applying Theorem III.1 on (IV.5), we conclude that the sequence $\{\mathbb{E}[X(t)]\}$ reaches consensus. Hence, the deterministic sequence $\{\Pi^{\perp}\mathbb{E}[X(t)] = \mathbb{E}[\Pi^{\perp}X(t)]\}$ is stable by Theorem II.1, i.e., $\mathbb{E}[\Pi^{\perp}X(t)]\|_1 \to 0$. By Proposition IV.1, we can always assume $x \in (\mathbb{R}^+)^N$. Thus, $\Pi^{\perp}X(t) \in (\mathbb{R}^+)^N$, and this leads to

$$\mathbb{E}[\|\Pi^{\perp}X(t)\|_\infty] \leq \mathbb{E}[\|\Pi^{\perp}X(t)\|_1] = \|\mathbb{E}[\Pi^{\perp}X(t)]\|_1 \to 0. \qquad \text{(IV.6)}$$

In other words, $\{\Pi^{\perp}X(t)\}$ is stable in L^1. This implies the consensus of μ by Theorem II.1.

■

Remark IV.1. In (IV.6), we used the fact, for all $Y : \Omega \to (\mathbb{R}^+)^N$,

$$\mathbb{E}[\|Y\|_1] = \mathbb{E}\Big[\sum_{i=1}^{N} Y_i\Big] = \sum_{i=1}^{N}\mathbb{E}[Y_i] = \|\mathbb{E}Y\|_1.$$

However, one shall not expect identity $\mathbb{E}[\|Y\|_\infty] = \|\mathbb{E}[Y]\|_\infty$ holds in general. For instance, one has strict a inequality, if $Y(\omega_1) = (0,1)^T$, $Y(\omega_2) = (1,0)^T$, and $\mathbb{P}(\{\omega_1\}) = \mathbb{P}(\{\omega_2\}) = 1/2$, then $\mathbb{E}[\|Y\|_\infty] = 1 > \frac{1}{2} = \|\mathbb{E}[Y]\|_\infty$ holds. This is the reason that we use the l^1-norm in (IV.6) instead of directly but incorrectly using $\mathbb{E}[\|\Pi^{\perp}X(t)\|_\infty] = \|\mathbb{E}[\Pi^{\perp}X(t)]\|_\infty \to 0$.

Corollary IV.1. *Consider* $\{X(t)\}$ *defined in* (II.3) *generated by distribution* μ. μ *reaches consensus if and only if* $\mathbb{E}^{\mu}[A(1)]$ *reaches consensus.*

V. Concluding remarks

In this paper, we derived a necessary and sufficient condition for consensus over a linear random network based on the connection between consensus and stability as shown by Theorem II.1. Although our proof is shown under the discrete-time framework, the similar results still hold in the continuous-time setting. Consequently, one can similarly follow the procedure to

obtain consensus conditions based on stability results of [5] on hybrid switching continuous-time systems.

Regarding the second-order random network, one can also utilize the result of this work. More precisely, for

$$X(t) = \alpha A(t)X(t-1) + \beta B(t)X(t-2), \quad t = 2, 3, \ldots$$

where $\alpha + \beta = 1$, $\alpha, \beta \geq 0$, $\{A(t)\}$ and $\{B(t)\}$ are i.i.d. stochastic matrix sequence with given distributions μ_A and μ_B on S_N, the problem is equivalent to

$$Y(t) = C(t)Y(t-1),$$

where $Y(t) = \begin{bmatrix} X(t) \\ X(t-1) \end{bmatrix}$ is an $\mathbb{R}^{2N}$-vector and $C(t) = \begin{bmatrix} \alpha A(t) & \beta B(t) \\ I & 0 \end{bmatrix}$ is a stochastic matrix in S_{2N}.

One last interesting remark is the application of Kolmogorov's 0-1 law [2], which is firstly given in [4] in the context of ergodicity of i.i.d. matrix sequence. Similar result also holds in the consensus and the stability problems. For instance, now we know that $X^x(t)$ defined in (II.3) does not reach consensus when $\lambda_2(\mathbb{E}[A(1)]) \geq 1$ for a given distribution μ in any of the three modes. In other words,

$$\mathbb{P}\Big\{\omega \in \Omega : X^x(t,\omega) \text{ reaches consensus for all } x \in \mathbb{R}^N\Big\} < 1.$$

Natural question is then, what the above probability is when $\lambda_2(\mathbb{E}[A(1)]) \geq 1$. The answer is surprisingly simple: zero.

Proposition V.1. *Consider $X^x(t)$ defined in* (II.3) *generated by i.i.d. matrices with distribution μ. Then*

$$\mathbb{P}\Big\{\omega \in \Omega : X^x(t,\omega) \textit{ reaches consensus for all } x \in \mathbb{R}^N\Big\}$$

is either 1 or 0.

Proof can be accomplished similarly to [4, Lemma 1], by using the tail σ-field argument on decreasing events of the form

$$B_k = \Big\{\omega : \Pi_{t=k}^{\infty} A(t)x \text{ reaches consensus for all } x \in \mathbb{R}^N\Big\}.$$

REFERENCES

[1] A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original edition.

[2] R. Durrett. *Probability: Theory and examples*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 3rd edition, 2005.

[3] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Trans. Automat. Control*, 50(11):1867–1872, 2005.

[4] A. Tahbaz-Salehi and A. Jadbabaie. A necessary and sufficient condition for consensus over random networks. *IEEE Trans. Automat. Control*, 53(3):791–795, 2008.

[5] C. Zhu, G. Yin, and Q. S. Song. Stability of random-switching systems of differential equations. *Quart. Appl. Math.*, 67(2):201–220, 2009.